Hypothesis Testing (Chapter 8)

Given a random sample $X_1, X_2, \ldots, X_n$, $X_i = f(X(i))$, we want to test if

$$H_0: \theta \in \Theta_0 \quad \text{or} \quad H_1: \theta \in \Theta_1$$

where $\Theta_0$ and $\Theta_1$ are subsets of the parameter space.

Example: Suppose that the average life of bulbs made under a standard manufacturing procedure is 1,400 hours. Now, suppose that a new method of sealing is designed and we wish to determine if this new method increases the life of the bulbs.

If $X_i$ = lifetime of bulb $i$, $i = 1, 2, \ldots, n$ and $X_i \sim \exp(\theta)$

$$f(x|\theta) = \frac{1}{\theta} \exp\left(-\frac{x}{\theta}\right)$$

we wish to test

$$H_0: \theta = 1400 \text{ hrs} \quad \text{or} \quad H_1: \theta > 1400 \text{ hrs}$$

In general, a hypothesis is a statement about the population parameter $\theta$.

The two hypotheses in a hypothesis testing are called null ($H_0$) and alternative ($H_1$).

We will restrict to situations with only two hypotheses.

Examples: In a seed classification program, we wanted to find if the level of germination was "poor", "intermediate" and "high".

If $p$ is the percentage of seeds that germinate over all the population:

- "poor" $\equiv p \leq 0.4$
- "intermediate" $\equiv 0.4 \leq p < 0.95$
- "high" $\equiv p \geq 0.95$
One variable of interest is:

$$X = \begin{cases} 
0 & \text{"no germination"} \\
1 & \text{"germination"} 
\end{cases}$$

$$X_1, X_2, \ldots, X_n \text{ iid. } X \sim \text{Bernoulli}(p)$$

$$H_0: 0 < p < 0.9; \quad H_1: p \geq 0.95; \quad H_2: p < 0.95.$$

Frequentist approach may or should with a loss-optimality of a Bayesian approach.

Back to two hypothesis only.

Usually, $$\Theta_1 = \Theta_0$$ and $$\Theta_0 \cap \Theta_1 = \emptyset$$

We also will classify hypothesis as "simple" or "composite".

Simple hypothesis: The distribution of $$X$$ is fully specified. (e.g., light bulb: $$\Theta = 1400$$ is simple $$\Rightarrow X \sim \text{Exp}(1400)$$).

Composite hypothesis: Distribution of $$X$$ not completely specified. (e.g., $$\Theta > 1400$$) simple vs. alternative.

Hypothesis testing procedure: is a rule that specifies that sample values $$x$$ for which $$H_0$$ is true, and for which sample values, $$H_1$$ is considered true.

The procedure is defined by a set $$\Theta$$ contained in $$\Theta$$ such that:

- If the sample point $$x \in \Theta$$ we reject $$H_0$$.
- If $$x \notin \Theta$$ we accept $$H_0$$.

Example: Suppose that $$X_1, X_2, \ldots, X_n$$ follow a $$N(\mu, 1)$$

$$H_0: \mu = 5 \text{ vs. } H_1: \mu > 5$$

We know $$T = \bar{x}$$ is a "good" estimator for $$\mu$$. 
A possible rejection region is:
\[ G = \{ x \in X : \sum_{i=1}^{n} x_i \geq 5 \}. \]
How to find \( G \)? How to evaluate \( G \)?

Methods of finding tests: Likelihood Ratio Test

Recall that under random sampling:
\[ L(\theta|x) = f(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta) \quad \text{for } H_0: \theta \in \Theta_0 \text{ vs. } H_1: \theta \in \Theta_1. \]

The likelihood ratio test is based on the likelihood ratio statistic:
\[ \lambda(x) = \sup_{\Theta_0} L(\theta|x) / \sup_{\Theta} L(\theta|x) \]

A likelihood ratio test is any test that has a rejection region of the form \( G = \{ x \in X : \lambda(x) \leq c \} \)
where \( c \) is any number such that \( 0 \leq c \leq 1 \).

Notice that: \( \lambda(x) \leq 1 \) (a restricted supremum is less than an unrestricted supremum).

Intuition: if \( H_0 \) is true, \( \sup_{\Theta_0} L(\theta|x) \approx \sup_{\Theta} L(\theta|x) \)
\( \Rightarrow \lambda(x) \approx 1 \). \( \Rightarrow \) Accept \( H_0 \) unless relative to \( \sup_{\Theta} L(\theta|x) \)

If \( H_0 \) is not true, \( \sup_{\Theta_0} L(\theta|x) \approx \sup_{\Theta} L(\theta|x) \Rightarrow \lambda(x) \leq c \)
\( \Rightarrow \) Reject \( H_0 \)

If \( \hat{\theta} \) is the MLE under \( H_0 \) and \( \hat{\theta} \) is the general MLE, then
\[ \lambda(x) = \frac{L(\hat{\theta}_0|x)}{L(\hat{\theta}|x)}. \]

Example: let \( x_1, x_2, \ldots, x_n \) be a random sample from the Poisson distribution:
\[ f(x|\theta) = \frac{\theta^x e^{-\theta}}{x!}; \quad x = 0, 1, 2, 3, \ldots. \]
Suppose \( H_0 : \theta = \theta_0 \) and \( H_1 : \theta \neq \theta_0 \) where \( \theta_0 \) is specified by the experimenter.

Find a critical region given by the likelihood ratio test.

The likelihood function is:

\[
L(\theta|x) = \frac{n! \theta^{\sum_i x_i} e^{-n\theta}}{\prod_i x_i!} = \frac{\theta^{\sum_i x_i} e^{-n\theta}}{\prod_i x_i!}
\]

\[
\sup L(\theta|x) = L(\theta_0|x) = \theta_0^{\sum_i x_i} e^{-n\theta_0} / \prod_i x_i!
\]

For the MLE:

\[
\log L(\theta|x) = \sum_i x_i \log \theta - n\theta - \log \prod_i x_i!
\]

\[
\frac{d \log L(\theta|x)}{d\theta} = \sum_i x_i / \theta - n = 0 \iff \hat{\theta} = \bar{x}
\]

\[
\sup L(\theta|x) = (\bar{x})^{\sum_i x_i} e^{-n\bar{x}} / \prod_i x_i!
\]

\[
\lambda(x) = \left( \frac{\theta_0^{\sum_i x_i} e^{-n\theta_0}}{\prod_i x_i!} \right) / \left( \frac{\hat{x}^{\sum_i x_i} e^{-n\bar{x}}}{\prod_i x_i!} \right)
\]

\[
= \left( \frac{\theta_0^{\sum_i x_i} e^{-n\theta_0}}{\hat{x}^{\sum_i x_i} e^{-n\bar{x}}} \right) = e^{-n(\theta_0 - \bar{x})} e^{n(\theta_0 - \bar{x})}
\]

The rejection region is defined by:

\[
E \left( \frac{\theta_0}{\bar{x}} \right) e < c
\]

Obs: The rejection region depends on \( \bar{x} \), a sufficient statistic for \( \theta \).

If \( T(x) \) is a sufficient statistic for \( \theta \) by the factorization theorem:

\[
f(x|\theta) = g(T(x)|\theta) h(x)
\]

\[
\lambda(x) = \frac{\sup L(\theta|x)}{\sup L(\theta|x)} = \frac{\sup g(T(x)|\theta) h(x)}{\sup g(T(\theta_0)|\theta) h(x)} = \frac{h(x) \sup g(T(x)|\theta)}{h(x) \sup g(T(\theta_0)|\theta)}
\]
THE LIKELIHOOD RATIO TEST TAKES A SIMPLER FORM WHEN

\[ \theta = \frac{\Theta}{0.01} \]

Ho: \( \theta = \Theta \) vs. Hi: \( \theta = \Theta_1 \)

\[ \sup \{ L(\theta|x) = L(\theta_0|x) \} \quad \sup \{ L(\theta|x) = \begin{cases} L(\theta_0|x) & \text{if } L(\theta_0|x) > L(\theta_1|x) \\ L(\theta_1|x) & \text{if } L(\theta_0|x) \leq L(\theta_1|x) \end{cases} \} \]

\[ \Rightarrow \Lambda(x) = \begin{cases} 1 & \text{if } L(\theta_0|x) > L(\theta_1|x) \\ \frac{L(\theta_0|x)}{L(\theta_1|x)} & \text{if } L(\theta_0|x) \leq L(\theta_1|x) \end{cases} \]

EXAMPLE: Suppose \( x_1, x_2, \ldots, x_n \) are iid observations with \( X \sim \text{Bernoulli}(\theta) \)

Ho: \( \theta = 0.5 \) vs. Hi: \( \theta = 0.9 \)

\[ f(x|\theta) = \theta^x (1-\theta)^{1-x}; \quad x=0,1; \quad L(\theta|x) = \theta^x (1-\theta)^{1-x}; \quad \theta_0 = 0.5, \theta_1 = 0.9. \]

\( L(x) \) will be determined by the LIKELIHOOD RATIO:

\[ \frac{L(\theta=0.5|x)}{L(\theta=0.9|x)} = \frac{(0.5)^{\sum x_i} (0.5)^n}{(0.9)^{\sum x_i} (0.1)^n} = \frac{(0.5)^{\sum x_i}}{(0.9)^{\sum x_i}} \]

\[ \Rightarrow \text{Reject Ho when} \]

\[ \frac{(0.5)^{\sum x_i}}{(0.9)^{\sum x_i}} < c; \quad \text{for } 0 < c < 1 \]

We express the inequality as:

\[ \text{Reject Ho } \iff \sum_{i=1}^{n} x_i > \frac{\log c + n \log (0.1/0.5)}{\log (0.1/0.9)} \]
Another example of LRT:

Suppose now, we want to test: \( x_1, x_2, \ldots, x_n \sim \text{Bernoulli}(\theta) \)

\( H_0: \theta \leq 0.5 \) vs. \( H_1: \theta > 0.5 \)

\[ L(\theta | x) = \begin{cases} L(\hat{\theta} | x); & \text{if } \hat{\theta} < 0.5 \rightarrow \text{NLE} \\ L(\theta = 0.5 | x); & \text{if } \hat{\theta} \geq 0.5 \end{cases} \]

\[ \hat{\theta} \]

\[ L(\theta | x) = L(\hat{\theta} | x) \Rightarrow \]

\[ \chi(x) = \begin{cases} 1 & \text{if } x < 0.5 \quad \text{(do not reject } H_0) \\ \frac{L(\theta = 0.5 | x)}{L(\hat{\theta} | x)} & x \geq 0.5 \end{cases} \]

\[ L(\theta = 0.5 | x) = \frac{(0.5)^{\sum x_i} (0.5)^{n-\sum x_i}}{(\sum x_i)^{\sum x_i} (1-\sum x_i)^n} = \left( \frac{\text{deg} (1-x)}{x} \right)^{\sum x_i} \left( \frac{0.5}{(1-x)} \right)^n \]

Reject \( H_0 \) if \( \left( \frac{1-x}{x} \right)^{\sum x_i} \left( \frac{0.5}{(1-x)} \right)^n < C \).

How to determine \( C \)? We’ll see this in the next section.

Bayesian Tests.

Recall that in Bayesian, we have a prior distribution \( \pi(\theta) \) which leads to the posterior distribution \( \pi(\theta | x) \).

\( H_0: \theta \in \Theta_0 \) vs. \( H_1: \theta \in \Theta_1 \)

We compute the posterior probabilities \( P(\theta \in \Theta_0 | x) \) and \( P(\theta \in \Theta_1 | x) \)

and compare them.

If \( P(\theta \in \Theta_0 | x) \geq P(\theta \in \Theta_1 | x) \Rightarrow \text{Accept } H_0 \)

If \( P(\theta \in \Theta_1 | x) > P(\theta \in \Theta_0 | x) \Rightarrow \text{Reject } H_0 \)
IN FACT

\[ P(\Theta \in \Theta_0 | x) = \int_{\Theta_0} \pi(\Theta | x) d\Theta \]

\[ P(\Theta \in \Theta_1 | x) = \int_{\Theta_1} \pi(\Theta | x) d\Theta \]

If \( \Theta_1 = \Theta_0 \) \( \Rightarrow \) \( P(\Theta \in \Theta_1) = 1 - P(\Theta \in \Theta_0) \)

**Accepting** \( \Theta_0 \) if \( P(\Theta \in \Theta_0) > \frac{1}{2} \)

**Problem:**

If \( \Theta_0 : \{ \Theta \in \Theta_0 \} \Rightarrow H_0 : \Theta = \theta_0 \) and \( H_1 : \Theta \neq \theta_0 \).

With \( \pi(\Theta | x) \) a **continuous** distribution \( \Rightarrow \)

\[ P(\Theta = \theta_0 | x) = 0 \] for any sample \( x \)

We will always reject \( H_0 \)! With **simple hypothesis**, the **Bayesian approach** has some **difficulties**.

**On the other hand**, if we have more than 2 hypothesis, for example,

\( H_0 : \Theta \in \Theta_0 \) ; \( H_1 : \Theta \in \Theta_1 \) ; \( H_2 : \Theta \in \Theta_2 \)

We compute

\[ P(\Theta \in \Theta_0 | x) ; P(\Theta \in \Theta_1 | x) ; P(\Theta \in \Theta_2 | x) \]

and select the hypothesis with the highest posterior probability.

**Example:** Let \( x_1, x_2 \). \( X_n \) i.i.d observations with a **Bernoulli distribution** of parameter \( \Theta \)

\( \pi(\Theta) \sim U(0, 1) \). We wish to test:

\( H_0 : \Theta \leq 0.5 \ vs. \ H_1 : \Theta > 0.5 \)

Since a \( U(0, 1) \) is a **Beta** (1, 1) distribution

The posterior distribution is a **Beta** \( (n', n+1) \)
where \( d' = 1 + \sum_{i=1}^{n} x_i \); \( \rho' = 1 + n - \sum_{i=1}^{n} x_i \)

\[ P(\theta \leq 0.5 | x) = P(\theta \leq 0.5 | x) = \int_{0}^{0.5} \frac{\Gamma(n+\alpha)}{\Gamma(\alpha) \Gamma(n+\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} d\theta \]

**Cumulative Beta Distribution**

**Establishes**

If our data suggests that \( n = 100 \), \( \sum x_i = 40 \)

**The posterior is a Beta** (41, 61)

**Under this Beta:**

\[ P(\theta \leq 0.5 | x) = 0.976978 \]

\[ P(\theta > 0.5 | x) = 1 - 0.976978 = 0.023022 \]

**Then, we accept Ho.**

If \( \sum x_i = 70 \) then the posterior is a Beta (31, 31)

And

\[ P(\theta \leq 0.5 | x) = 0.00000274 \]

**Alternatively, we accept Ho if**

\[ \frac{1}{2} \leq P(\theta \leq 0.5 | x). \]

\[ \Rightarrow m \leq 0.5 \; ; \; \text{where} \; m \; \text{is the median of a Beta} \; (d', \rho'). \]

**Reject Ho otherwise**

In the first case, \( m = 0.40131 < 0.5 \Rightarrow \text{Reject Ho.} \)

In the second case, \( m = 0.6973641 \geq 0.5 \Rightarrow \text{Accept Ho.} \)
METHODS OF EVALUATING TESTS

Error Probabilities and the Power Function

The following table illustrates the types of errors in hypothesis testing.

<table>
<thead>
<tr>
<th>Decision</th>
<th>Accept $H_0$</th>
<th>Reject $H_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$</td>
<td>Correct</td>
<td>Type I Error</td>
</tr>
<tr>
<td>$H_1$</td>
<td>Type II Error</td>
<td>Correct</td>
</tr>
</tbody>
</table>

Type I Error $\equiv$ Reject $H_0$ when $H_0$ is true.

Type II Error $\equiv$ Accept $H_0$ when $H_0$ is false.

Key Probabilities:

If $C$ (or $R$) is the rejection region and we wish to test: $H_0: \theta \in \Theta_0$ vs. $H_1: \theta \in \Theta_1$.

The probability of a Type I Error is:

$Pr(\theta \in \Theta_1 | \theta \in \Theta_0)$ where $\theta \in \Theta_0$.

The probability of a Type II Error is:

$Pr(\theta \notin \Theta_1 | \theta \in \Theta_0)$ where $\theta \in \Theta_0$.

Both types of errors depend on $Pr(\theta \in \Theta_1 | \theta \in \Theta_0)$.

This leads to the following definitions:

Def: The power function of a hypothesis test with rejection regions $C$ is a function of $\theta$ defined by

$\beta(\theta) = Pr(\theta \in \Theta_1 | \theta \in \Theta_0)$

The ideal power function is:

$\beta(\theta) = \begin{cases} 0 & \text{if } \theta \in \Theta_0 \\ 1 & \text{if } \theta \in \Theta_1 \end{cases}$
Usually, this ideal cannot be attained.

A "good" power function is such that:

\[ \beta(\theta) \approx 1 \text{ if } \theta \approx \theta_0 \text{ and } \beta(\theta) \approx 0 \text{ if } \theta \approx \Theta_0. \]

Example: Let \( X_1, X_2. \) \( X_n \) be a random sample from a normal \((\theta, 25)\). Consider

\( H_0: \theta \leq 17 \) vs. \( H_1: \theta > 17. \)

Reject \( H_0 \) if and only if \( \bar{X} > 17 + \frac{5}{\sqrt{n}} \) (This a
LRT see exercise 8.37.)

Find the power function \( \beta(\theta) \)

\[ \beta(\theta) = \Pr \left[ \bar{X} > 17 + \frac{5}{\sqrt{n}} \mid \theta \right]. \text{ Notice that given } \theta, \bar{X} \sim N(\theta, 25/n) \Rightarrow Z = \frac{\bar{X} - \theta}{\frac{5}{\sqrt{n}}} \]

Then

\[ \beta(\theta) = \Pr \left( Z > \frac{17 + \frac{5}{\sqrt{n}} - \theta}{\frac{5}{\sqrt{n}}} \right) = 1 - \Phi \left( \frac{17 + \frac{5}{\sqrt{n}} - \theta}{\frac{5}{\sqrt{n}}} \right) \]

Where \( \Phi \) is the cdf of a \( N(0,1) \).

As \( \theta \) increases from \(-\infty\) to \( \infty \) \( \Phi \left( \frac{17 + \frac{5}{\sqrt{n}} - \theta}{\frac{5}{\sqrt{n}}} \right) \) decreases from 1 to 0, then \( \beta(\theta) \) increases from 0 to 1.

Sketch of \( \beta(\theta) \):

\[ \begin{array}{c}
\theta \\downarrow \\
0.5 \downarrow \\
17 + \frac{5}{\sqrt{n}} \downarrow \\
0 \uparrow \\
\theta \uparrow \\
\end{array} \]

\( \theta < 17 \) power \( \approx 0 \) \( \checkmark \) This is what we want

\( \theta > \theta_1 \) power \( \approx 1 \) \( \checkmark \)

For \( 17 < \theta < \frac{5}{\sqrt{n}} + A \) power \( < 0.5 \) less than \( 1/2 \) of rejecting \( H_0 \).
NOTICE THAT THE MAXIMUM TYPE I ERROR IS REACHED AT $\theta = \theta_0 = 17$

WHAT IS THIS MAXIMUM TYPE I ERROR PROBABILITY?

$\beta(\theta=17) = 1 - \Phi(1) = 0.1586$

WHAT IS THE MAXIMUM TYPE II ERROR IF $\theta = 17 + 5$?

RECALL THAT IF $\theta \in \Theta_1$,

$\Pr[\text{Type II Error}] = 1 - \beta(\theta)$

THE MAXIMUM TYPE II ERROR FOR $\theta = 17 + 5$ IS

$1 - \beta(\theta = 22) = \Phi \left( \frac{5/\sqrt{n} - 5}{5/\sqrt{n}} \right)$ DEPENDS ON $n$!

IF $n$ FIXED (SAY $n = 25$) $\Rightarrow 1 - \beta(\theta = 22) = \Phi(-4) \approx 0$

IF WE FIX THIS ERROR (SAY AT $0.01$) WE CAN FIND $n$ THAT YIELDS THIS DESIRED PROBABILITY.

$\Rightarrow \Phi \left( \frac{5/\sqrt{n} - 5}{5/\sqrt{n}} \right) = 0.01$

$\Rightarrow \frac{5/\sqrt{n} - 5}{5/\sqrt{n}} = -2.33$ SOLVE FOR $n$

$n \approx 11.0889 \Rightarrow n = 12$ GIVES THE DESIRED PROBABILITY.

GLOBALLY TYPED TYPE II ERROR THE ERRORS WILL DEPEND STRONGLY ON $n$.

THE CRITICAL SIZE OF A TEST IS $\alpha = sup \beta(\theta)$ $\theta \in \Theta_0$.
A test with power $\beta(\theta)$ is of level $\alpha$ if

$$\sup_{\theta \in \Theta} \beta(\theta) \leq \alpha$$

In complicated testing situations, it is often impossible to build a size $\alpha$ test.

A test with power $\beta(\theta)$ is of size $\alpha$ if

$$\sup_{\theta \in \Theta} \beta(\theta) = \alpha$$

In one example, $\alpha = 0.1586$ is the size of test.

Back to LRT,

Reject $H_0 \iff \lambda(x) \leq C$

To determine $C$, first we specify a level $\alpha$ (say 0.01, 0.05).

The power function is:

$$\beta(\theta) = Pr \left[ \lambda(x) \leq C \mid \theta \right]$$

$c$ is such that:

$$\sup_{\theta \in \Theta} Pr \left[ \lambda(x) \leq C \mid \theta \right] \leq \alpha$$

Example: Let $X_1, X_2, \ldots, X_n$ be iid Bernoulli$(\theta)$

$H_0: \theta = 0.5$ vs. $H_1: \theta = 0.9$

The LRT establishes that we reject $H_0$ if and only if:

$$\sum_{i=1}^{n} X_i > C$$

$C = 0.05$

The null hypothesis is formed of one point $\theta = 0.5$

$$\sup_{\theta \in \Theta} Pr \left( \sum_{i=1}^{n} X_i > C \mid \theta = 0.5 \right) = Pr \left( \sum_{i=1}^{n} X_i > C \right)$$

If $n$ is large enough, by CLT $\sum_{i=1}^{n} X_i \sim N(n(0.5), n(0.5))$

$$\Rightarrow Pr \left( \sum_{i=1}^{n} X_i > C \mid \theta = 0.5 \right) = Pr \left( Z > \frac{C - n(0.5)}{\sqrt{n(1-0.5)}} \right) = .05$$
\[ C = -1.64 \cdot \left( \frac{\ln(0.9)(0.1)}{n \cdot \ln(0.9)} \right) \]

Then solving for \( n \) and \( C \)

\[ n = 10.75 \quad ; \quad C = 8.0688 \]

(If assuming that the CLT is valid here, otherwise we need to use the CDF of a Binomial)

Most Powerful Test:

Suppose we wish to test

\[ H_0: \theta = \theta_0 \quad \text{vs.} \quad H_1: \theta = \theta_1 \]

A level \( \alpha \) test is the most powerful if its power function satisfies:

\[ \beta(\theta) \geq \beta'(\theta) \quad \text{for all} \quad \theta \in (\theta_0, \theta_1) \]

where \( \beta'(\theta) \) is the power function of any other \( \alpha \)-level test.

To easiest situation to find the most powerful test is when \( H_0: \theta = \theta_0 \quad \text{vs.} \quad H_1: \theta = \theta_1 \quad \theta_0 \neq \theta_1 \) (both hypotheses are simple). The power function is given by 2-point Neyman-Pearson Lemma. Consider testing \( H_0: \theta = \theta_0 \quad \text{vs.} \quad H_1: \theta = \theta_1 \) A test with rejection region \( R \) (or \( \bar{R} \)) such that

\[ x \in R \iff f(x | \theta_1) > k f(x | \theta_0) \]

\[ x \in \bar{R} \iff f(x | \theta_1) < k f(x | \theta_0) \quad \text{or} \quad k \]

And \( \alpha = \int_{x \in R} f(x | \theta_1) \) is a uniformly most powerful test (UMP). \( \alpha \)-level test.
Proof only for the case where \( f(x|\theta) \) is a pdf of a continuous random variable.

Define \( \phi(x) \) by the test function defined by \( R \)
\[
\phi(x) = \begin{cases} 1 & \text{if } x \in R \\ 0 & \text{if } x \notin R. \end{cases}
\]

Let \( \phi'(x) \) be the test function of any other level test \((\phi'(x) = 1; x \in R'; \phi'(x) = 0, \text{if } x \notin R').\)

Let \( \beta(\theta) \) be the power function of \( \phi(x) \) and \( \beta'(\theta) \) be the power function of \( \phi'(x) \).

Since \( 0 \leq \phi(x) \leq 1 \), we have \( \phi(x) = 1 \) if \( f(x|\theta_0) > K f(x|\theta) \) and \( \phi = 0 \) if \( f(x|\theta_1) < K f(x|\theta) \)
\[
\Rightarrow \quad [\phi(x) - \phi'(x)] [f(x|\theta_1) - K f(x|\theta_0)] \geq 0
\]
\[
\Rightarrow \quad 0 \leq \int [\phi(x) - \phi'(x)] [f(x|\theta_1) - K f(x|\theta_0)] \, dx
\]
\[
= \int \phi(x) f(x|\theta_1) \, dx - \int \phi'(x) f(x|\theta_1) \, dx - K \int \phi(x) f(x|\theta_0) \, dx
\]
\[
+ \int K \phi'(x) f(x|\theta_0) \, dx = \beta'(\theta_1) - \beta'(\theta_0) - K \beta(\theta_0) - K \beta'(\theta_0)
\]

We know that \( \beta(\theta_0) = 1 \) because \( \alpha = P(x \in R|\theta = \theta_0) \)
and \( \beta'(\theta_0) \leq 1 \) \( \Rightarrow \) 
\[
- K (\beta(\theta_0) - \beta'(\theta_0)) \geq 0
\]
\[
0 \leq \beta'(\theta_1) - \beta'(\theta_0) - K (\beta(\theta_0) - \beta'(\theta_0)) \leq \beta'(\theta_1) - \beta'(\theta_0)
\]
\[
\Rightarrow \quad \beta'(\theta_1) > \beta'(\theta_0) \Rightarrow R \text{ defines the}
\]

UMP.
\[
\Rightarrow \text{By the Normal CDF } \frac{C - n(0.5)}{\sqrt{n}(0.5)} = 1.64
\]

\[
\Rightarrow C = 1.64 \sqrt{n}(0.5) + n(0.5)
\]

IF CLT DOES NOT APPLY, RECALL THAT \(\sum_{i=1}^{n} X_i \sim Bin(n, \theta)\)

UNDER \(\theta_0 \Rightarrow Bin(n, 0.5)\), C HAS TO BE OBTAINED FROM THE BINOMIAL CDF. \((n = 10, \text{ C }= 7, \alpha = 0.05468\)

Observation: If \(C = 8, \alpha = 0.1074\)

\(H_0: \theta = \theta_0 \text{ vs. } H_1: \theta = \theta_1\)

AND \(\beta\) IS THE REJECTION REGION FOR TEST.

Error probabilities:

\[
\alpha = P(X \in R | \theta = \theta_0)
\]

\[
\beta = P(X \in R | \theta = \theta_1) = 1 - P(X \in R | \theta = \theta_1)
\]

In the previous, say \(n = 50, \alpha = 0.05\)

\[
\Rightarrow C = 1.64 \sqrt{50}(0.5) + 50(0.5)
\]

\[
= 30.80
\]

\(\text{What is } \beta?\)

\[
\beta = 1 - P \left( \frac{\sum_{i=1}^{n} X_i \geq 30.80}{\sqrt{50}(0.5)} | \theta = 0.9 \right)
\]

\[
= 1 - P \left( \frac{Z \geq 30.80 - 50(0.9)}{\sqrt{50}(0.5)} \right) = 1 - P(Z > 6.7) \approx 0
\]

IF \(\alpha, \beta\) ARE DETERMINED, WE CAN FIND \(n\) THAT LEADS TO THESE RECOMMENDATIONS. RECALL BULL EX. \(H_0: \theta = 0.5 \text{ vs. } H_1: \theta = 0.9\)

Suppose we want that \(\alpha = \beta = 0.05\) for the bull ex.

Reject \(H_0\) IF \(\sum X_i > 7C\)

\[
P(\sum X_i > 7C | \theta = 0.5) = 0.05 \Rightarrow C = 1.64 \sqrt{n}(0.5) + n(0.5)
\]

\[
0.05 = P(\sum X_i > 7C | \theta = 0.9) \Rightarrow \frac{C - n(0.9)}{\sqrt{n(0.9)(0.1)}} = -1.64
\]
---

**Notes:**

- **Revers of Neyman-Pearson Lemma is True.**
  - If $\phi'$ is a most powerful test of size $\alpha$, necessarily $\phi' = \phi$; $\phi$ test function for Neyman-Pearson.
  - (See end of Theorem 8.3.12).

- The **N-P Lemma Says:** $\text{Reject } H_0 \iff k f(x|100) < f(x|101)$
  - Which is the same to
    \[ \text{Reject } H_0 \iff \frac{f(x|100)}{f(x|101)} < \frac{1 - c(L.R.T.)}{K} \]

**Example:**

Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x; \theta) = \theta e^{-\theta x}$ where $\theta = \theta_0$ or $\theta = \theta_1$; $\theta_0$ and $\theta_1$ are known fixed numbers. Assume that $\theta_1 > \theta_0$.

Now,

\[ f(x|100) = \theta_0^n \exp(-\theta_0 \sum x_i); \quad f(x|101) = \theta_1^n \exp(-\theta_1 \sum x_i) \]

According to the Neyman-Pearson Lemma, the most powerful test has the form:

\[ \text{Reject } H_0 \iff k < \left( \frac{\theta_1}{\theta_0} \right)^n \exp\left( - \left( \theta_1 - \theta_0 \right) \frac{\sum x_i}{\theta_1} \right) \]

which is equivalent to

\[ \sum_{i=1}^{n} x_i < \left( \frac{1}{\theta_1 - \theta_0} \right) \log \left( \frac{\theta_1^n}{k} \right) = k' \]

The inequality of the N-P Lemma has been simplified to:

\[ \sum_{i=1}^{n} x_i < k' \]
Also, \( \alpha = P_{\theta} \left[ \text{reject } H_0 \mid \theta = \theta_0 \right] = P_{\theta} \left[ \sum_{i=1}^{n} x_i < k' \mid \theta = \theta_0 \right] \)

We know that \( \sum_{i=1}^{n} x_i \sim \text{Gamma}(n, \theta^{-1}) \).

Hence,
\[
P_{\theta} \left[ \sum_{i=1}^{n} x_i < k' \mid \theta = \theta_0 \right] = \int_{-\infty}^{k'} \frac{\theta_0^n}{\Gamma(n)} x^{n-1} e^{-\theta_0 x} \, dx = \alpha
\]

\( k' \) is the \( \alpha \) - quantile of a Gamma \( (n, \theta^{-1}) \).

How to find a UMP if at least the alternative is composite.

Example: Let \( X_1, X_2 \). \( X_i \) be iid such that \( X_i \sim N(\mu, \sigma^2) \) with \( \sigma^2 \) known. The hypothesis to test are

\( H_0: \mu = \mu_0 \) vs. \( H_1: \mu > \mu_0 \)

Idea: Fix a value and perform a simple vs. simple test. Take an arbitrary \( \mu \) in the alternative.

Find N-P rejection region.

\[ \text{Reject } H_0 \iff \frac{f(x \mid \mu_0)}{f(x \mid \mu_1)} \leq \frac{1}{k} \]

\[
f(x \mid \mu_0) = \left( \frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu_0)^2 \right)
\]

\[
f(x \mid \mu_1) = \left( \frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu_1)^2 \right)
\]

Recall that \( \sum_{i=1}^{n} (x_i - \mu)^2 = n(\bar{x} - \mu)^2 + \sum_{i=1}^{n} (x_i - \bar{x})^2 \)

\[ \Rightarrow \frac{f(x \mid \mu_0)}{f(x \mid \mu_1)} = \exp \left( -\frac{n}{2\sigma^2} \left[ (\bar{x} - \mu_0)^2 - (\bar{x} - \mu_1)^2 \right] \right) \leq \frac{1}{k} \]

After some algebra...
The inequality is equivalent to
\[
\frac{n}{\sigma^2} (\bar{X} - \mu_0) X < \log (\frac{1}{\alpha}) + \frac{n}{2 \sigma^2} (\mu_0^2 - \mu_1^2)
\]

Since \((\mu_0 - \mu_1) < 0\) then our region has the form
\[
\bar{X} > k'
\]

This form does not depend on \(\mu_1\) and \(k'\) is fixed with \(\alpha = \text{Pr} [\bar{X} > k' | \mu = \mu_0] = (k' = \mu_0 + z_{\alpha/2})\)

Hence, the region \(\bar{X} > k'\) defines the most pow. test.

What happens when we consider the test?
\[
H_0: \mu = \mu_0 \text{ vs. } H_1: \mu < \mu_0
\]

And let's say we use the same rejection region.
\[
R = \{ \bar{X} > k' \}; \quad d = \text{Pr} [\bar{X} > k' | \mu = \mu_0]
\]

Is this a \(\alpha\)-level UMP?

All we need to check is that:
\[
d = \sup_{\Theta} \beta(d)
\]

Where
\[
\beta(d) = \text{Pr} [\bar{X} > k' | \mu] = \text{Pr} [z > \frac{k' - \mu}{\sigma / \sqrt{n}}]
\]

\[
\beta(d) = 1 - \Phi (\frac{k' - \mu}{\sigma / \sqrt{n}})
\]

What if we now want to test
\[
H_0: \mu = \mu_0 \text{ vs. } H_1: \mu < \mu_0 \quad (k' = \mu_0 - z_{\alpha/2} \frac{\sigma}{\sqrt{n}})
\]
The rejection region is now of the form
\[
\bar{X} < k' \text{ this is also}
\]
\[
N = \text{Pr} [\bar{X} < k' | \mu = \mu_0]
\]

A UMP.
WHAT ABOUT $H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$?

For $\mu < \mu_0$, the most powerful test is given by $R_1$:

Reject $X < k_1$ with a power function $\beta_1(\mu) = P[X < k_1 | \mu] \uparrow$

For $\mu > \mu_0$, the most powerful test is given by $X > k_2$ with a power function $\beta_2(\mu) = P[X > k_2 | \mu] \downarrow$

At $\mu_1$

$\beta_1(\mu_1) \uparrow \beta_2(\mu_1)$

At $\mu_2$

$\beta_2(\mu_2) \uparrow \beta_1(\mu_2)$

There is not a UMP test for this problem.

TRADEOFF:

Reject $H_0$ if $\bar{X} < k''$ or $\bar{X} > k'$.

(Two tailed test)

$\alpha = P \left\{ \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < \frac{-2\sqrt{2}}{n} \right\}$

$\Rightarrow k'' = \mu_0 - 2\sqrt{2} \left( \frac{\sigma}{\sqrt{n}} \right)$

$\Rightarrow k' = \mu_0 + 2\sqrt{2} \left( \frac{\sigma}{\sqrt{n}} \right)$

This is not a UMP test at level $\alpha$.
LAST TIME: \( x_1, x_2, \ldots, x_n \) iid obs. \( X \sim N(\mu, \sigma^2) \) \( \sigma^2 \) known.

\[ H_0: \mu = \mu_0 \text{ vs. } H_1: \mu \neq \mu_0 \]

If we use as our rejection region \( \Phi = \frac{1}{\sqrt{2\pi}} \frac{x - \mu_0}{\sigma/\sqrt{n}} > z_{\alpha/2} \)

or \( \frac{x - \mu_0}{\sigma/\sqrt{n}} < -z_{\alpha/2} \) this leads to a size \( \alpha \) TEST but not a UMP TEST.

AnotherAngel. \( \Phi \) is a L.R.T. !

\[ L(\mu; x) = \left( \frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right) \]

\[ \sup_{H_0} L(\mu; x) = \left( \frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (\bar{x} - \mu_0)^2 \right) \]

\[ \sup_{\mu} L(\mu; x) = \left( \frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right) \]

The RATIO IS:

\[ \lambda(x) = \frac{\exp \left( -\frac{1}{2\sigma^2} \left( n(\mu_0 - \bar{x})^2 + \sum_{i=1}^{n} (x_i - \bar{x})^2 \right) \right)}{\exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)} = \exp \left( \frac{-n(\mu_0 - \bar{x})^2}{2\sigma^2} \right) \]

Reject \( H_0 \) if \( \lambda(x) \leq c \Leftrightarrow \exp \left( -\frac{n}{2\sigma^2} (\bar{x} - \mu_0)^2 \right) \leq c \)

\[ \frac{(\bar{x} - \mu_0)^2}{\sigma^2/n} \geq \frac{-2 \log(c)}{\sigma^2/n} \rightarrow \frac{1}{\sigma^2/n} \geq \frac{-2 \log(c)}{\sigma^2/n} \]

\[ (-2 \log(c))^{1/2} \text{ FIXED TO } z_{\alpha/2} \]

WITH \( \sigma^2 \) UNKNOWN.

UNDER \( H_0 \):

THE MLE FOR \( \sigma^2 \) IS \( \hat{\sigma}_{\text{MLE}}^2 = \frac{n}{n-2} \sum_{i=1}^{n} (x_i - \bar{x})^2 \)

Recall that the unrestricted MLE were:

\( \hat{\mu} = \bar{x} \) AND \( \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \)
\[
\sup L(\mu, \sigma^2 | x) = \left( \frac{1}{2\pi \sigma^2} \right)^{n/2} \exp \left( -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right)
\]

\[
\sup L(\mu, \sigma^2 | x) = \left( \frac{1}{2\pi \sigma^2} \right)^{n/2} \exp \left( -\frac{n}{2\sigma^2} \right)
\]

\[
\lambda(x) = \frac{\sup L(\mu, \sigma^2 | x)}{\sup L(\mu, \sigma^2 | x)} = \left( \frac{\sigma^2}{\hat{\sigma}^2} \right)^{n/2} \text{ Reject } H_0 \iff \left( \frac{\sigma^2}{\hat{\sigma}^2} \right)^{n/2} \leq C
\]

Recall that:
\[
\sum_{i=1}^{n} (x_i - \mu)^2 = n (\bar{x} - \mu_0)^2 + \sum_{i=1}^{n} (x_i - \bar{x})^2
\]

Reject \( H_0 \iff \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n (\bar{x} - \mu_0)^2} \geq C \]

Den. of \( \lambda(x) \)

Reject \( H_0 \iff c^{n/2} \leq \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n (\bar{x} - \mu_0)^2} \) if we multiply by \( (n-1) \)

Reject \( H_0 \iff \frac{|\bar{x} - \mu_0|}{S/\sqrt{n}} > k \) where \( S^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 \)

Given \( H_0 \) follows a \( t(n-1) \) \( \Rightarrow k = t(n-1) \)

Note: By a similar argument to \( \chi^2 \) known case this is not a UMP.

In general for \( H_0: \theta = \theta_0 \) vs \( H_1: \theta \neq \theta_0 \) a UMP will exist (or \( \theta < \theta_0 \))

For \( H_0: \theta = \theta_0 \) vs. \( H_1: \theta \neq \theta_0 \) usually a UMP will not exist.
A large class of problems that have an UMP d-level test for one sided hypotheses involve the nonmonotone likelihood ratio property.

What is a nonmonotone likelihood ratio?

Suppose we have a statistic $T$ with a family of pdfs or pmfs $g(t; \theta); \theta \in \Theta$ of a real parameter $\theta$. It has a MLR if for every $i, j \in \Theta$, the ratio $g(t; \theta_i)/g(t; \theta_j)$ is a monotonic (inc. or dec.) function of $t$.

Example: If $X_1, X_2, \ldots, X_n$ are i.i.d. Poisson($\lambda$) RVs, we know $T = \sum X_i \sim \text{Poisson}(n\lambda)$. $g(t; \lambda) = \frac{e^{-\lambda \lambda^t}}{t!}; t = 0, 1, 2, \ldots$. Consider $\lambda_1 > \lambda_2$.

$g(t; \lambda_1)/g(t; \lambda_2) = e^{-\lambda_1 \lambda_2} \left( \frac{\lambda_1}{\lambda_2} \right)^t$ increasing in $T$.

If $X_1, X_2, X_3, \ldots, X_n$ are i.i.d. Bernoulli($\theta$), we have

$T = \sum X_i \sim \text{Bin}(n; \theta), g(t; \theta) = \binom{n}{t} \theta^t (1-\theta)^{n-t}$

If $\theta_1 > \theta_2 \Rightarrow g(t; \theta_1)/g(t; \theta_2) = \left( \frac{\theta_1}{\theta_2} \right)^t \left( \frac{1-\theta_2}{1-\theta_1} \right)^{n-t}$ increasing in $t$.

In fact, $g(t; \theta) = h(t) \phi(\theta) e^{w(\theta)t} \text{ (Exp. Family)}$ has a MLR if $w(\theta)$ is a non-decreasing function.

If $\theta_1 > \theta_2 \Rightarrow g(t; \theta_1)/g(t; \theta_2) = \left( \frac{w(\theta_1)}{w(\theta_2)} \right)^t e^{(w(\theta_1)-w(\theta_2))t}$

Since $w(\cdot)$ non-decreasing $\Rightarrow$ ratio non-decreasing in $t$. The MLR looks beyond the exponential family.
EXAMPLE: Let $X_1, X_2, \ldots, X_n$, i.i.d. $\sim U(0, \theta)$

\[
T = \max\{X_1, X_2, \ldots, X_n\} \quad g(t|\theta) = \frac{n t^{n-1}}{\theta^n}, \quad 0 < t < \theta
\]

\[
p_{\theta_1 > \theta_2} \Rightarrow \frac{g(t|\theta_1)}{g(t|\theta_2)} = \left(\frac{\theta_2}{\theta_1}\right)^n \frac{I(t)/I(t)}{(\theta_1, \theta_1) / (\theta_2, \theta_2)}
\]

RATIO only DEFINED FOR $g(t|\theta_0)$, $g(t|\theta_0)$.

MAIN RESULT FOR MLE. (KARLIN-RUBIN THEOREM 8.3.17)

$H_0: \theta \leq \theta_0$ vs. $H_1: \theta > \theta_0$

$T$ IS A SUFFICIENT STATISTIC FOR $\theta$ WITH FAMILY

$\{g(t|\theta); \theta \in \Theta\}$ WITH A MLE. THEN THE RULE

Reject $H_0 \iff T > t_0$; where $\alpha = P(T > t_0 \mid \Theta_0)$

DEFINES A UMP OF SIZE $\alpha$.

$H_0: \theta = \theta_0$ vs. $H_1: \theta > \theta_0 \Rightarrow T > t_0$ DEFINES THE REJECTION REGION.

1) $H_0: \theta \leq \theta_0$ vs. $H_1: \theta > \theta_0$ REJECT $H_0 \iff \sum_{i=1}^{n} X_i > t_0$

$\alpha = P\left[\sum_{i=1}^{n} X_i > t_0 \mid \sum_{i=1}^{n} \sim \text{Poisson}(n \lambda_0)\right]$

2) $H_0: \theta = \theta_0$ vs. $H_1: \theta > \theta_0$ REJECT $H_0 \iff \sum_{i=1}^{n} X_i > t_0$

$\alpha = P\left[\sum_{i=1}^{n} X_i > t_0 \mid \sum_{i=1}^{n} \sim \text{Bin}(n, \theta)\right]$

3) $H_0: \theta \leq \theta_0$ vs. $H_1: \theta > \theta_0$ REJECT $H_0 \iff X_{(n)} > t_0$

$\alpha = P\left[X_{(n)} > t_0 \mid \theta = \theta_0\right] = \int_0^{t_0} \frac{b_n^{-n} dt}{\theta^n} = \int_0^{t_0} \frac{e^{-b_n t}}{\theta^n} dt$

$\alpha = \left(\frac{\theta_0^n - t_0^n}{\theta_0^n}\right) \Rightarrow t_0 = \left(\theta_0^n - \alpha \theta_0^n\right)^{1/n}$